Effective Omnistereo Panorama Video Generation by Deformable Spheres

Sheng-Kuen Huang
National Taiwan University
victorhsk1114@gmail.com

Hong-Shiang Lin
National Taiwan University
amsdya@gmail.com

Ming Ouhyoung
National Taiwan University
ming@csie.ntu.edu.tw

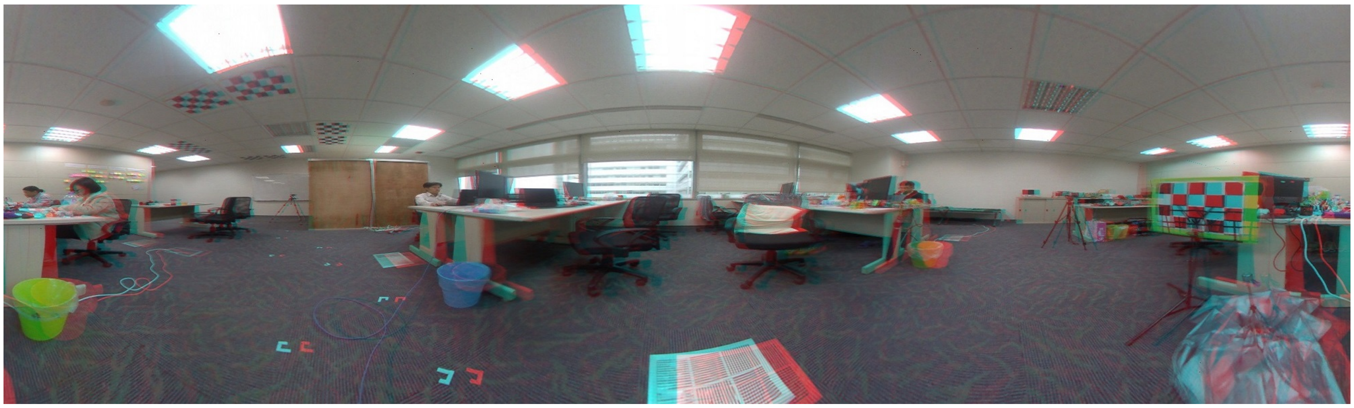

Figure 1: One frame of an omnistereo panorama generated from our method, shown in anaglyphic stereo here.

ABSTRACT

We presents an alternative and effective method for omnistereo panorama video generation by using deformable spheres. Deformable spheres use vertex based spherical mesh to represent the depth of a scene. Meanwhile, the spherical mesh was substituted for the original 3D points cloud mesh in the rendering step. In our experiments, this approach reduces the time of omnistereo panorama generation and the result is smooth in both spatial and temporal domain.

CCS CONCEPTS

• **Computing methodologies → Image and video acquisition**; *Computational photography*;

KEYWORDS

Omnistereo panorama, Video stitching , Spherical representation

ACM Reference format:
Sheng-Kuen Huang, Hong-Shiang Lin, and Ming Ouhyoung. 2017. Effective Omnistereo Panorama Video Generation by Deformable Spheres. In *Proceedings of SIGGRAPH '17 Posters, Los Angeles, CA, USA, July 30 - August 03, 2017,* 3 pages.
https://doi.org/10.1145/3102163.3102199

SIGGRAPH '17 Posters, July 30 - August 03, 2017, Los Angeles, CA, USA

ACM ISBN 978-1-4503-5015-0/17/07.
https://doi.org/10.1145/3102163.3102199

1 INTRODUCTION

Virtual reality (VR) is one of the major trends over the world, and there are lots of reserch topics about it. One of the topics is about omnistereo panorama/ omnidirectional stereo (ODS) video generation, such as Google jump[Anderson et al. 2016]. An omnistereo panorama video has a pair panoramas video, where each frame has one panorama video for the left eye and the other for the right eye. Two panorama can provide depth for human and make a user to have more immersed experience.

In our previous work, we proposed a method use image based rendering to generate high quality omnistereo panorama. We used stereo matching to reconstruct 3D points cloud and calculate virtual view points to simulate the human eye position. Then, we used ray tracing to generate the final ODS. However, the computation time of previous work is big in both steps and there exsit some artifacts which was produced by pixel-wise ray tracing.

In this work, we propose to use deformable spheres, which is proposed by Lee et al[Lee et al. 2016], to represented depth of the scene. Deformable sphere is a vertex based spherical mesh where each vertexes can change its radius to fit the scene, and at the same time to insure the temporal coherence and surface smoothness by adding smooth constrain. The proposed approach can greatly reduce the time cost in depth estimating and rendering step. Another advantage in using deformable sphere is to have less artifacts by rendering. In our experiment, we use different vertex number and compared the time cost with previous method.

2 OUR APPROACH

We first estimate depth map by deformable spheres for our system. Our system consists of 4 pre-calibrated 180-degree fish-eye cameras and each camera is located at 90-degree in rotation. Then,we

	previous method	10x10	20x20	40x40	100x100
(a)	over 500	0.157	0.736	6.624	103.863
(b)	743.98	47.72	62.36	66.47	70.85

Table 1: Average computation time per frame in seconds. (a) is depth estimation step. (b) is rendering step.

do feature detection and matching for adjacent cameras in the overlapping area. For each feature pair, we use camera parameters to calculate real position of this pair in the 3D space with the centroid of 4 cameras as the origin. After construct the feature 3D point cloud, then use the technique of Lee[Lee et al. 2016]. By minimizing the energy function, we try to insure that the surface of sphere mesh is as smooth as possible in first and second differential and also can fit the 3D points cloud which are constructed by features. In the meantime, we also insure the deformable sphere also have consistency in temporal space by adding temporal constraints. In rendering step, we replace the origin 3D point cloud by our deformable sphere mesh and calculate viewing circle image which was proposed by Peleg et al[Peleg et al. 2001].

In our experiments, we observe that our approach can greatly increase the speed in both the depth estimation step and the rendering step. Since our mesh is vertex based, our time complexity can be 400-500% smaller as compared to the origin pixel based algorithm. Since deformable sphere can use different vertex numbers, we experimented with vertex numbers at 10*10, 20*20, 40*40 and 100*100.

Table 1 shows the average computation time for different vertex number when solve the linear system in depth estimated step and rendering step. Obviously,more vertexes need more time to construct matrix and solving. In our experiment, the computation time is no more than 500 seconds.If use per-pixel local stereo matching method, the cost time will more than 500 seconds. In rendering step, we render 2K size omnistereo panoramas and our method speed up the computation time to 10-12 times.

We observe that with more vertex number it is more flexible to fit the scene, but the panorama also need more time to render and solve linear system, as shown in Figure 2.

Our second advantage is our results are more smooth than original. Compared to pixel-based rendering, our result doesn't have lots of artifact dot since we use a mesh to replace all the 3D points cloud. Using vertex-based mesh to render just like warping, so the result is more clearer than previous algorithm,as shown in Figure 3.

3 CONCLUSION AND FUTURE WORK

Our approach can increase the efficiency of the omnistereo panorama video generation. Secondly, our result is also smooth on spatial and temporal domain.

Although we can efficiently generate omnistereo panorama video, the result is still sensitive to the deformable sphere mesh. If we don't have enough feature points or too few vertexes to estimate the depth, the panorama can still have strange artifacts. In future work, we may use adaptive mesh to make deformable sphere more flexible to fit any scene. Since our algorithm also use rendering technique, the GPU version software is undergoing. This may also accelerate the speed of our algorithm.

Figure 2: Comparison of different vertex numbers and their pair of result.

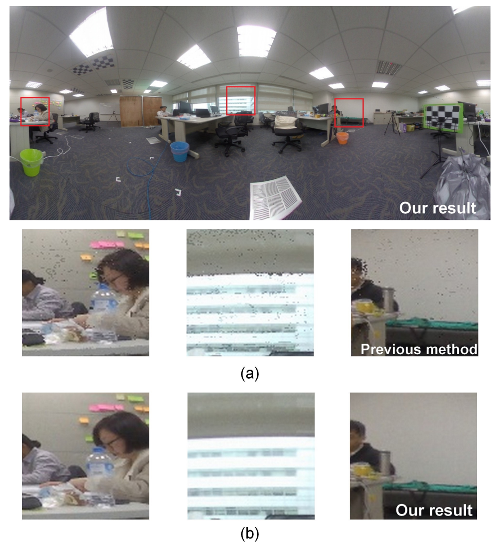

Figure 3: Compared our method to previous method. Top: Our result use 40*40 vertex numbers. (a)Part of our previous results and have noise as artifacts. (b) Improvement of our results in this work.

4 ACKNOWLEDGEMENTS

This work was partially supported by MOST 105-2622-8-002-002 (MediaTek Inc.) and MOST 103-2221-E-002-171-MY2.

REFERENCES

Robert Anderson, David Gallup, Jonathan T. Barron, Janne Kontkanen, Noah Snavely, Carlos Hernandez Esteban, Sameer Agarwal, and Steven M. Seitz. 2016. Jump: Virtual Reality Video.

Jungjin Lee, Bumki Kim, Kyehyun Kim, Younghui Kim, and Junyong Noh. 2016. Rich360: Optimized Spherical Representation from Structured Panoramic Camera Arrays. *ACM Trans. Graph.* 35, 4, Article 63 (July 2016), 11 pages. https://doi.org/10.1145/2897824.2925983

S. Peleg, M. Ben-Ezra, and Y. Pritch. 2001. Omnistereo: panoramic stereo imaging. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23, 3 (Mar 2001), 279–290. https://doi.org/10.1109/34.910880